

KOMBINATORIKUS – KOMPOZICIONÁLIS REPREZENTÁCIÓK

GYENES VIKTOR

TÉZISFÜZET PH.D. DISSZERTÁCIÓHOZ

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

INFORMATIKAI KAR

INFORMÁCIÓS RENDSZEREK TANSZÉK

Témavezető: Prof. habil. Lőrincz András

Informatikai Doktori Iskola

Dr. Demetrovics János

Információs Rendszerek Program

Dr. Benczúr András

Budapest, 2011.

Bevezetés

A gépi tanulás a számítástudomány egyik leggyorsabban növekvő ágazata, azzal a nagyratörő céllal, hogy az emberi gondolkodást modellezze. Azonban viszonylag fiatal terület lévén sok részlet még nem teljesen tisztázott. A gépi tanulási módszerek gyakran alkalmaznak elosztott reprezentációt, ami azt jelenti, hogy a bemeneti adatokat vagy állapotokat *komponensekre bontják*. Ezen komponensek pontos definíciója módszerenként más és más; különböző feladatok különböző természetű komponenseket igényelnek építőelemekként. Néhány módszer célja, hogy megtalálja egy adott feladat reprezentációjához szükséges komponenseket – más néven *jellemzőket* – és újra *egyesítse* őket az érvelés, a felhasználás során.

A kompozicionalitás talán legismertebb példája a nyelv; építőelemei a szavak, amiket mondatokká kombinálhatunk. A nyelvben megjelenő kompozicionalitás feltehetően a gondolkodásunkban, a mentális reprezentációinkban jelenlévő kompozicionalitást tükrözi. Emiatt természetes, hogy a számítástudomány sok ágazata alkalmaz kompozicionális reprezentációkat. De hogyan tanulják meg az emberek egy adott feladat reprezentációját? Például, hogyan tanulja meg egy sakkjátékos, hogy hova fókuszálja a figyelmét a sakktáblán? Gyakran, amikor az emberi gondolkodást magas szinten modellezzük, szabályrendszert használunk: ‘*ha* valamilyen feltétel fennáll *akkor* az valamilyen akciót vagy értéket von maga után’. A feltétel általában egyszeű állítások *konjunkciója*, míg konjunkciók *diszjunkciója* külön szabályként realizálható. A sakkban egy ilyen szabály lehetne például: ‘*ha* az ellenfél le tudja ütni egy bábumat és én nem tudok cserébe leütni az övéi közül egyet *akkor* ez előnytelen számomra’.

Ezen disszertáció a fenti szabályokhoz hasonló komponenseket használ a feladatokkal kapcsolatos tudás reprezentációjára, melynek alapjául a feladatot leíró változók *kombinációi* szolgálnak. Ezen munka célja az ily módon előálló komponensek és a belőlük felépülő reprezentáció tulajdonságainak és használhatóságának vizsgálata.

A dolgozat során a *kombinatorikus* és a *kompozicionális* szó két ellentétes, mégis szorosan kapcsolódó fogalmat jelöl. A *kompozicionalitás* általában valami olyasmit sugall, amit fel lehet építeni, vagy ki lehet következtetni a komponenseiből. Például, egy mondat jelentése többé kevésbé következik a szavainak jelentéséből, habár a szavak sorrendje és nyelvtani kapcsolata is befolyásolhatja a pontos jelentést.

Ezzel ellentétben, ha egy feladat *kombinatorikus*, az általában azt jelenti, hogy efféle dekompozíció nem alkalmazható. A matematikában például a kombinatorikus optimalizáció olyan feladatokkal foglalkozik, amik során különböző állapotok kerülnek kiértékelésre, melyek változók kombinációval írhatók le, és előfordulhat, hogy az összes lehetséges kombinációt ki kell értékelni az optimum megtalálásához, mivel kissé különböző kombinációknak nagyon különböző értéke lehet.

Azonban e dolgozat mellett érvel, hogy ez a két ellentétes fogalom jól kiegészíti egymást. Tömören összefoglalva, *ha egy feladatból kivonjuk a benne rejlő kombinatorikusságot, akkor ami marad, az kompozicionális*. Visszatérve a nyelvi példához, ha megtanuljuk, hogy mely szó-kombinációknak van speciális jelentése, a maradék már kikövetkeztethető a szavak jelentéséből, azaz a kompozicionalitás segítségével.

Módszerek

A kompozicionalitással kapcsolatos vizsgálataim kiinduló pontja a nyelv, melynek során a kompozicionális szókincs fejlődését modellezem úgynevezett *nyelvi játékok* segítségével. E nyelvi játékok egyik alaptétele az, hogy a résztvevők kompozicionális állapotleírással rendelkeznek, és ezen állapotukról kívánnak kommunikálni. A következő lépés, amit a dolgozat elemez, a kompozicionális reprezentációk alkalmazhatósága a döntéshozás modellezésben. E célra a megerősítéssel tanulás szigorú keretrendszere nyújt alapot, annak is a kompozicionális kiterjesztése, az úgynevezett *faktorizált megerősítéssel tanulás*. Végül, inspirációt merítve a faktorizált megerősítéssel tanulási feladatokban felmerülő struktúra keresésből, általános módszereket vezetek be függvények kompozicionális közelítésére, nagy hangsúlyt fektetve a megfelelő komponensek keresésére. A következő fejezet röviden ismerteti a dolgozatban kifejlesztett algoritmusok és módszerek leírására használt matematikai modellt.

Alapvető Matematikai Modell

A matematikában a kompozicionalitás modellje gyakran a lineáris függvény: komponens függvények súlyozott összege. A fenti gondolatok formalizálására használt modell alapja: állapotváltozók direkt-szorzata fölött definiált függvények közelítése változó-

kombinációkra épülő komponens-függvények segítségével. Ha a bevezetőben említett gondolatmenetet a kombinatorikusság kivonásáról erre a területre vonatkoztatjuk, a következőt kapjuk: ha megtaláljuk mely állapotváltozó-kombinációkhoz tartoznak kitüntetett értékek, akkor a függvényt közelíthetjük ezen kombinációk – a kivonatolt komponensek – értékeinek összegeként. Ez a filozófia húzódik a dolgozatban kifejlesztett alapvető tanulási architektúra és módszerek mögött.

Legyen $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_d$ egy d változóból álló, a feladatot leíró direkt-szorzat állapottér. A legtöbb gépi tanulási módszer direkt vagy indirekt módon épít egy vagy több közelítendő $f : \mathcal{X} \rightarrow \mathbb{R}$ függvényre. Például a felügyelt tanulási feladatok, mint például a regresszió, vagy az osztályozás explicit módon próbálnak tanulni egy hozzárendelést, míg a megerősítéses tanulás optimális viselkedést tanul megerősítéseken keresztül indirekt módon, azonban a megoldási módszerek valamilyen hozzárendelést tartanak fenn, mint például állapot-átmeneteket, állapot-értékeléseket. Azaz, a legtöbb esetben a feladat lényegében *függvény-approximációt* rejt magában. Az approximáció legegyszerűbb formájában a függvényt lineáris formában keressük:

$$\hat{f}(x) = \sum_{i=1}^n \mathbf{w}_i \phi_i(x) = \mathbf{w}^T \phi(x) ,$$

ahol $x \in \mathcal{X}$, és $\phi : \mathcal{X} \rightarrow \mathbb{R}^n$ egy n komponensből, más néven jellemzőkből álló függvény, és $\mathbf{w} \in \mathbb{R}^n$ a megtanulandó paraméterek (súlyok) vektora. Ha a jellemzők halmaza kellően gazdag, tetszőleges függvény jól közelíthető a fenti architektúrával. Ezen dolgozatban használt architektúra specialitása a jellemzők definíciójában rejlik.

Mint azt már a bevezetőben említettük, az itt használt komponensek *állapotváltozó-kombinációk* formájában állnak elő. Formálisan, legyen $I \subset \{1, \dots, d\}$ egy tetszőleges index halmaz. Jelölje $|\mathcal{X}_i|$ a (diszkrét) \mathcal{X}_i változó által felvehető értékek számát, és tételizzük fel az általánosság megszorítása nélkül, hogy ezek az $\{1, \dots, |\mathcal{X}_i|\}$ értékek. Legyen $\mathbf{k} \in \mathbb{N}^{|I|}$ egy egész értékű vektor, $1 \leq \mathbf{k}_j \leq |\mathcal{X}_{I_j}|$ minden $j = 1, \dots, |I|$, azaz \mathbf{k}_j jelölje az \mathcal{X}_{I_j} változó által éppen felvett értéket. Ekkor

$$\phi_{\mathbf{k}}(x) := \delta \left(\bigwedge_{j=1}^{|I|} \mathcal{X}_{I_j} = \mathbf{k}_j \right)$$

egy *kombinációs jellemzője* az x állapotnak, ahol δ a Kronecker függvényt jelöli, amelynek értéke 1 ha az argumentuma igaz, és 0 ha hamis.

A kombinációs jellemzők általánosíthatók folytonos változókra fuzzy reprezentáció segítségével, melynek során az állapotváltozók konjunkciója a fuzzy bázisértékekhez tartozó valószínűségek *szorzatával* helyettesíthető. Egy másik lehetőség az általánosításra maguknak a (folytonos, valós) állapotváltozóknak a szorzata; így *monomokat* kapunk jellemzőként, a teljes architektúra pedig egy többváltozós közelítő *polinom*má válik.

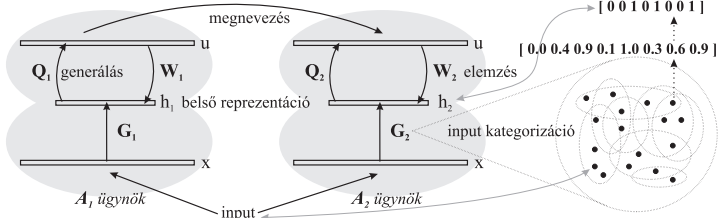
A fent vázolt, kombinációs jellemzőkre épített lineáris függvény-approximáció modellekre úgy is tekinthetünk, mint egy *neurális hálóra* egy kombinációs vagy szorzat jellemzőkből álló belső réteggel.

Tézisek

Eredményeim három területre sorolhatók, az előző fejezetben bevezetett metodológiából adódóan: a kompozicionális nyelv kialakulásának modellezése, kompozicionális reprezentációk használatának vizsgálata a döntéshozás modellezésben, és algoritmusok fejlesztése jellemzők generálására függvény-approximáció során.

Nyelvi kompozicionalitás

A kompozicionális nyelvi modellezés során kifejlesztettem egy ügynök architektúrát és tanulási algoritmust amelynek középpontjában az ügynökök kompozicionális állapotrepresentációja áll, a nyelvi játék paradigmába [10] ágyazva. E megközelítés során, az ügynökök nyelvi elemekkel illetik az általuk megfigyelt világot, és a cél egy közös szókincs kialakítása a megfigyelések jelölésére. Az architektúra két fontos új gondolatot vezet be a kompozicionalitás kapcsán: az úgynevezett *rekonstrukciós elv* használatát kompozicionális reprezentációk esetén, illetve egy egyszerű modellt a kompozicionális reprezentáció hangolására az együttes tanulás során. A rekonstrukciós elv egy módszer a nyelvi megnyilvánulások generálására és elemzésére, amelyet egy rekonstrukciós neurális háló valósít meg: egy nyelvi megnyilvánulást úgy elemez, hogy az olyan belső reprezentációt eredményezzen, amely leginkább az adott nyelvi megnyilvánulásra ad okot (azaz leginkább rekonstruálja azt).



Ügynök architektúra kompozicionális nyelvi játékhöz. Két ügynök, A_1 és A_2 struktúrája. A belső reprezentáció (h) az inputok (x) kategorizációjából áll elő (G transzformáció), amit valószínűségek paramétereznek és különbözhetnek a két ügynökben. A Q és W transzformációk a nyelvi elemek generálását és elemzését írják le, a rekonstrukciós elv segítségével. Az ügynökök a játék során hangolják a Q , W és G transzformációkat, miközben egyikük ‘megnevez’ (u) egy inputot a másik pedig ‘figyel’.

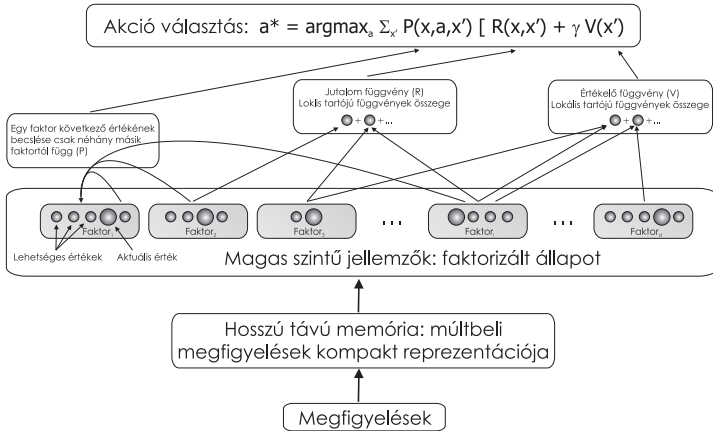
Ezen kívül vizsgáltam a több ügynök együttes tanulásával járó instabilitási problémát, amely abból fakad, hogy az ügynökök tanulási folyamatai egymástól függenek.

Tézis 1. A kompozicionális nyelv előnyei együttes szókincs kialakítása során kompozicionális állapotreprezentációra alapozva

- Megmutattam, hogy az együttes tanulás akár egyszerű kommunikációs körülmények között is divergálhat amennyiben a kommunikáció költséggel jár. Ezen kívül megmutattam, hogy ha az ügynökök modellezzik egymást, a probléma súlyossága csökkenthető [4]. A rekonstrukciós elv egyfajta nyelvi modell egymásra nézve
- Szimulációkkal demonstráltam, hogy amennyiben a rekonstrukciós elvet használják, az ügynökök egy konzisztens szókincsben állapodnak meg, míg ez a megállapodás gyakran nem alakul ki a rekonstrukciós elv használata nélkül [2] [3]
- Szimulációkkal demonstráltam, hogy a kompozicionális szókincs kialakulása nagyságrendekkel jobban skáláz az állapotleírás méretével és a résztvevő ügynökök számával mint a holisztikus szókincsé [2] [3]
- Szimulációkkal demonstráltam, hogy ha az inputok kategorizációja különbözik az ügynökökben, kompozicionális nyelv esetén lehetséges azok összehangolása és hatékonyabb kommunikáció elérése, de holisztikus nyelv esetén nem [2] [3]

Kompozicionalitás a Megerősítéses Tanulásban

A megerősítéses tanulás [11] egy a Markov döntési folyamatok (MDP) köré épített keretrendszer szekvenciális döntéshozatal modellezésére. Sok kapcsolódó tanulási algoritmus központi eleme egy *értékelő függvény*, amit az ügynök az állapotok és akciók kiértékeléséhez tanul. Alapvető algoritmusok léteznek arra az esetre, amikor az ügynök minden állapotot külön értékkel lát el. Ez a megközelítés nem alkalmazható nagy állapotterek, realisztikus feladatok esetén. Azonban, ha az állapottér változók direkt-szorozataként definiált, és az állapotváltozók közötti függőségek gyengék, tömörebb feladatleírásra és megoldási módszerekre nyílik lehetőség. Az így kapott keretrendszert faktorizált megerősítéses tanulásnak nevezzük [6]. Ebben az esetben az értékelő függvény faktorizált formában közelíthető, ami, mint ahogy dolgozatomban megmutatom, a fent definiált kombinációs jellemzőkön alapuló lineáris formával ekvivalens.



Ügynök architektúra faktorizált megerősítéses tanuláshoz. A Markovi állapottér megfigyelésekből, hosszú távú memória segítségével kivonott állapotváltozókból épül fel. Az állapotátmeneti (P) jutalom (R) és értékelő (V) függvények lokális tartójú függvények kompozíciói. Az akció választás ezen függvények alapján történik.

A megoldási módszerek egy népszerű és a gyakorlatban nagyon hasznos ágazata szimulált tapasztalatra épül, amely esetén az ügynök inkrementálisan tanul a világgal

történő interakciója során. Ezen módszerek legjelentősebb képviselője az úgynevezett *időbeli differenciák módszere*. Habár sok faktorizált megerősítéses tanulási módszert találunk az irodalomban, az időbeli differenciák módszerét úgy tűnik hanyagolták.

Tézis 2. A faktorizált időbeli differenciák módszerének konvergenciája

- *Illusztráltam a kapcsolatot a kombinációs jellemzők és az állapotér-particionálás között, és megmutattam, hogy a közelítő tanulás faktorizált MDP-ben strukturálisan kapcsolódik a pontos tanuláshoz egy nem-faktorizált segéd MDP-ben*
- *Bizonyítást adtam több inkrementális közelítő módszer konvergenciájára faktorizált MDP-ben, beleértve a faktorizált időbeli differenciák módszerét. Ezen eredmények központi eleme az értékelő függvény lineáris közelítése*
- *Szimulációkkal demonstráltam, hogy a faktorizált időbeli differenciák módszere exponenciálisan gyorsabb lehet mint a hagyományos, táblázatos változat [1]*
- *Szimulációkkal demonstráltam, hogy egy bizonyos rekurrens neurális hálózat (Echo State Network [9]), ami egy időben kiterjesztett lineáris approximátor, alkalmas megerősítéses tanulásra nem-Markovi környezetekben [5]*

Kompozicionalitás a Függvény-Approximációban

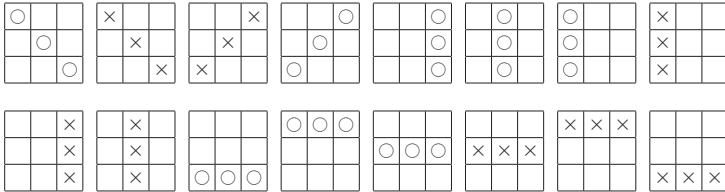
Inspirációt merítve a kombinációs jellemzők használatából a faktorizált megerősítéses tanulásban, általánosságban foglalkoztam a kombinációs jellemzők kiválasztásával függvények közelítése céljából. A faktorizált MDP-ben is használt, *regressziós fa építésen* [7] alapuló struktúra tanulási technikákból kiindulva, az ismert módszereket kombinációs jellemzők generálására adaptáltam lineáris függvény approximáció során. A lineáris forma révén a módszer továbbfejleszthetővé vált a *lépésenkénti regresszió* (más néven ortogonális legkisebb négyzetes közelítés [8]) egy inkrementális változatává. Az inkrementalitás ebben az esetben azt jelenti, hogy a kiválasztható jellemzők nem ismertek előre, hanem a folyamat közben, a kombinációk komplexitása szerint növekvő sorrendben generáljuk őket.

Algoritmusok két fő ágazatát fejlesztettem ki: az egyik explicit módon épít *ortogonális* jellemzőkre (amely többváltozós Legendre polinomokkal történő közelítéshez

vezet folytonos esetben), a másik pedig implicit módon ortogonalizálja a jellemzőket a legkisebb négyzetes közelítés során. Az explicit ortogonalizáció egyszerűbb algoritmusokat eredményez, azonban csak speciális esetekben alkalmazható, amikor az állapot-tér egyenletesen mintavételezett, és minden állapotváltozó-kombináció megengedett.

Tézis 3. Algoritmusok kombinációs jellemzők inkrementális generálására diszkrét és folytonos függvények közelítése esetén

- Bizonyítottam, hogy diszkrét állapotváltozók esetén, ha minden változókombináció megengedett input, egy ortogonális jellemző mátrix explicit módon megadható. Folytonos esetben Legendre polinom jellemzőket kapunk
- Kifejlesztettem egy algoritmust, amely inkrementálisan generál és válogat ortogonális kombinációs jellemzőket függvény-approximáció során, és szimulációkkal demonstráltam a módszer hatékonyságát
- Kifejlesztettem egy algoritmust amely implicit ortogonalizációt alkalmazva inkrementálisan és mohó módon generálja, sorrendezi és ritkítja a kombinációs jellemzőket a közelítésben mért hasznosságuk alapján. Többféle példán demonstráltam, hogy nagyon pontos közelítések készíthetők alacsony számú jellemző használatával, sokkal jobbak mint regressziós fákkal



A legfontosabb kivonatolt jellemzők a Tic-Tac-Toe nevű játék jutalomfüggvényének közelítéséhez. A közelítés bemenete 27 bináris változó, cellánként három, amelyek azt jelzik, hogy az adott cella tartalmaz-e \times vagy \circ jelet, illetve, hogy üres-e. A fontos kombinációs jellemzők egy sorban, oszlopban vagy átlóban elhelyezkedő három egyforma szimbólumból épülnek fel, mint ahogy azt a játékszabályok is sugallják.

A Szerző Saját Publikációi

- [1] V. Gyenes, Á. Bontovics, and A. Lőrincz. Factored temporal difference learning in the New Ties environment. *Acta Cybernetica*, 18:651–668, 2008.
- [2] V. Gyenes and A. Lőrincz. Co-learning and the development of communication. In *Proceedings of the 17th International Conference on Artificial Neural Networks*, volume 4668 of *Lecture Notes in Computer Science I*, pages 827–837, 2007.
- [3] V. Gyenes and A. Lőrincz. Language development among co-learning agents. In *Proceedings of 6th International Conference on Development and Learning*, pages 294 – 299, 2007.
- [4] A. Lőrincz, V. Gyenes, M. Kiszlinger, and I. Szita. Mind model seems necessary for the emergence of communication. *Neural Information Processing - Letters and Reviews*, 11:109–121, 2007.
- [5] I. Szita, V. Gyenes, and A. Lőrincz. Reinforcement learning with echo state networks. In *Proceedings of the 16th International Conference on Artificial Neural Networks*, volume 4131 of *Lecture Notes in Computer Science II*, pages 830–839, 2006.

Külső Hivatkozások

- [6] C. Boutilier, R. Dearden, and M. Goldszmidt. Stochastic dynamic programming with factored representations. *Artificial Intelligence*, 121(1-2):49–107, 2000.
- [7] L. Breiman, J. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. Chapman & Hall, New York, NY, 1984.
- [8] S. Chen, S. Billings, and W. Luo. Orthogonal least squares methods and their application to non-linear system identification. *Internat. Journal of Control*, 50:1873–1896, 1989.
- [9] H. Jaeger and H. Haas. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless telecommunication. *Science*, 304(5667):78–80, 2004.
- [10] L. Steels. Grounding symbols through evolutionary language games. In *Simulating the Evolution of Language*, chapter 10, pages 211–226. Springer Verlag, 2002.
- [11] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.